

Amendments to the Claims:

This listing of claims will replace all prior versions, and listings, of claims in the application:

Listing of Claims:

1. (Currently Amended) A method of associating a phenotype with one or more candidate chromosomal regions in a genome of ~~an organism~~ a species using a phenotypic data structure that ~~represents~~ comprises a difference in a phenotype between different strains of said ~~organism~~ species, said genome including a plurality of loci, said method comprising:
 establishing a genotypic data structure, said genotypic data structure corresponding to a locus ~~selected from~~ in said plurality of loci, said genotypic data structure ~~representing~~ comprising a variation of at least one component of said locus between different strains of said species ~~organism~~;
 comparing said phenotypic data structure to said genotypic data structure to form a correlation value; and
 repeating said establishing and comparing steps for each locus in said plurality of loci, thereby identifying one or more genotypic data structures that form a high correlation value relative to all other genotypic data structures that are compared to said phenotypic data structure ~~during~~ by said comparing step; wherein the loci that correspond to said one or more genotypic data structures that form a high correlation value represent said one or more candidate chromosomal regions and wherein an amount of said genome that is included in each locus in said plurality of loci is predetermined.
2. (Cancelled)
3. (Currently amended) The method of claim 2 1, wherein said amount is selected from a value in the range of about 0.01 centiMorgans to about 100 centiMorgans.
4. (Currently amended) The method of claim 2 1, wherein said amount is selected from a value in the range of about 5 cM to about 30 cM.

5. (Original) The method of claim 1, wherein an instance of said establishing step comprises selecting a locus that is centered on a portion of said genome that is a predetermined distance away from the locus that was selected by a previous instance of said establishing step.
6. (Original) The method of claim 5, wherein said predetermined distance is measured in centiMorgans.
7. (Original) The method of claim 5, wherein said predetermined distance is selected from the range of about 0.0001 centiMorgans to about 30 centiMorgans.
8. (Original) The method of claim 5, wherein said predetermined distance is selected from the range of about 2 centiMorgans to about 15 centiMorgans.
9. (Currently amended) The method of claim 1, each element in said phenotypic data structure ~~representing~~ corresponding to a difference in a phenotype between different strains of said ~~species~~ organism; wherein, for each element in said phenotypic data structure, said different strains of said ~~organism~~ species are selected from a plurality of strains of said ~~organism~~ species.
10. (Currently amended) The method of claim 9, wherein said difference in said phenotype is determined by a measurement of an attribute corresponding to said phenotype in different strains of said ~~organism~~ species.
11. (Currently amended) The method of claim 1, each element in said phenotypic data structure ~~representing~~ corresponding to a difference in said phenotype between a first cluster of strains of said ~~organism~~ species and a different second cluster of strains of said ~~organism~~ species; wherein, for each element in said phenotypic data structure, said different first and second cluster of strains of said ~~organism~~ species are selected from a plurality of clusters of strains of said ~~organism~~ species.
12. (Currently amended) The method of claim 1, each element in said genotypic data structure ~~representing~~ corresponding to a ~~variation~~ difference of at least one component of said locus between different strains of said ~~organism~~ species; wherein, for each element in

said genotypic data structure, said different strains of said ~~organism~~ species are selected from a plurality of strains of said ~~organism~~ species.

13. (Currently amended) The method of claim 12, wherein an amount that a variation contributes to said at least one component of said locus between different strains of said ~~organism~~ species is a function of a distance said variation is away from a center of the locus that corresponds to said genotypic data structure.

14. (Currently amended) The method of claim 13, wherein said genotypic data structure ~~represents~~ comprises a plurality of variations that are distributed about the center of said locus, and said establishing step further comprises:

fitting a distribution of said plurality of variations about the center of said locus with a probability function; and

weighting each variation by a corresponding value derived from said probability function such that variations further from the center of said locus are downweighted so that they contribute less to said genotypic data structure than loci that are closer to said center of said locus.

15. (Original) The method of claim 14 wherein said probability function is a Gaussian probability distribution, a Poisson distribution, or a Lorentzian distribution.

16. (Currently amended) The method of claim 1, each element in said genotypic data structure ~~representing~~ comprising a variation of at least one component of said locus between a first cluster of strains of said ~~organism~~ species and a different second cluster of strains of said ~~organism~~ species; wherein, for each element in said genotypic data structure, said different first and second clusters of strains of said ~~organism~~ species are selected from a plurality of strains of said ~~organism~~ species.

17. (Original) The method of claim 1, wherein said correlation value is formed in accordance with the expression:

$$c(P, G^L) = \frac{\sum^i (p(i) - \langle P \rangle) (g(i) - \langle G^L \rangle)}{\{[\sum^i (p(i) - \langle P \rangle)^2] [\sum^i (g(i) - \langle G^L \rangle)^2]\}^{1/2}}$$

where,

$c(P, G^L)$ is said correlation value;
 $p(i)$ is a value of the i^{th} element of said phenotypic data structure;
 $g(i)$ is a value of the i^{th} element of said genotypic data structure;
 $\langle P \rangle$ is a mean value of all elements in said phenotypic data structure; and
 $\langle G^L \rangle$ is a mean value of all elements in said genotypic data structure.

18. (Original) The method of claim 1, wherein said correlation value is weighted by a number of components in said locus.

19. (Original) The method of claim 1, wherein each said component is a single nucleotide polymorphism.

20. (Currently amended) The method of claim 1, wherein said correlation value is formed in accordance with the expression:

$$c(P, G^L) = \frac{[\sum^i (p(i) - \langle P \rangle) (g(i) - \langle G^L \rangle)] \times Z}{\{[\sum^i (p(i) - \langle P \rangle)^2] [\sum^i (g(i) - \langle G^L \rangle)^2]\}^{1/2}}$$

where,

$c(P, G^L)$ is said correlation value;
 $p(i)$ is a value of the i^{th} element of said phenotypic data structure;
 $g(i)$ is a value of the i^{th} element of said genotypic data structure;
 $\langle P \rangle$ is a mean value of all elements in said phenotypic data structure;
 $\langle G^L \rangle$ is a mean value of all elements in said genotypic data structure; and
 Z is a function of a number of components in said locus having a variation between different ~~strands~~ strains of said ~~organism~~.

21. (Original) The method of claim 20, wherein said function is selected from the group consisting of taking the square root of Z , squaring Z , raising Z by the power of a positive integer, taking a logarithm of Z , and taking an exponential of Z .

22. (Currently amended) The method of claim 1, wherein said correlation value is a correlative measure cm that is computed in accordance with the expression:

$$cm(P, G^L) = \frac{[\sum^i (p(i) - \langle P \rangle) (g(i) - \langle G^L \rangle)]}{\{[\sum^i (p(i) - \langle P \rangle)^2]\}^{1/2}}$$

where,

$cm(P, G^L)$ is said correlative measure;

$p(i)$ is a value of the i^{th} element of said phenotypic data structure;

$g(i)$ is a value of the i^{th} element of said genotypic data structure;

$\langle P \rangle$ is a mean value of all elements in said phenotypic data structure; and

$\langle G^L \rangle$ is a mean value of all elements in said genotypic data structure;

23. (Original) The method of claim 1, wherein said correlation value is formed using an algorithm selected from the group consisting of regression analysis, regression analysis with data transformations, a Pearson correlation, a Spearman rank correlation, a regression tree and concomitant data reduction, partial least squares, and canonical analysis.

24. (Original) The method of claim 1, wherein said repeating step further comprises:

computing (i) a mean correlation value that represents a mean of each said correlation value formed during instances of said comparing step; and (ii) a standard deviation of said mean correlation value based on each said correlation value formed during instances of said comparing step;

wherein, said one or more genotypic data structures that form a high correlation value relative to all other genotypic data structures compared to said phenotypic data structure during said comparing step are identified by selecting genotypic data structures that form a correlation value that is a predetermined number of standard deviations above said mean correlation value.

25. (Original) The method of claim 1, wherein each said variation in said genotypic data structure is obtained from a variation in a single nucleotide polymorphism database, a microsatellite marker database, a restriction fragment length polymorphism database, a short tandem repeat database, a sequence length polymorphism database, or an expression profile database.

26. (Currently amended) A computer program product for use in conjunction with a computer system, the computer program product comprising a computer readable storage medium and a computer program mechanism embedded therein, the computer program mechanism comprising:

a genotypic database for storing variations in genomic sequences of a plurality of strains of ~~an organism~~ a species;

a phenotypic data structure ~~that represents~~, said phenotypic data structure comprising a difference in a phenotype between different strains of said ~~organism~~ species; and

a program module for associating a phenotype with one or more candidate chromosomal regions in a genome of said ~~organism~~ species, said genome including a plurality of loci, said program module comprising:

instructions for establishing a genotypic data structure, said genotypic data structure corresponding to a locus ~~selected from a~~ in said plurality of loci, said genotypic data structure ~~representing~~ comprising a variation of at least one component of said locus between different strains of said ~~organism~~ species stored in said genotypic database;

instructions for comparing said phenotypic data structure to said genotypic data structure to form a correlation value; and

instructions for repeating said instructions for establishing and said instructions for comparing, for each locus in said plurality of loci, thereby identifying one or more genotypic data structures that form a high correlation value relative to all other genotypic data structures that are compared to said phenotypic data structure by said instructions for comparing; wherein the loci that correspond to said one or more genotypic data structures that form a high correlation value represent said one or more candidate chromosomal regions and wherein an amount of said genome that is included in each locus in said plurality of loci is predetermined.

27. (Cancelled)

28. (Currently amended) The computer program product of claim ~~27~~ 26, wherein said amount is selected from a value in the range of about 0.01 centiMorgans to about 100 centiMorgans.

29. (Currently amended) The computer program product of claim ~~27~~ 26, wherein said amount is selected from a value in the range of about 5 cM to about 30 cM.

30. (Original) The computer program product of claim 26, wherein an instance of said instructions for establishing comprises instructions for selecting a locus that is centered on a portion of said genome that is a predetermined distance away from the locus that was selected by a previous instance of said instructions for establishing.

31. (Original) The computer program product of claim 30, wherein said predetermined distance is measured in centiMorgans.

32. (Original) The computer program product of claim 30, wherein said predetermined distance is selected from the range of about 0.0001 centiMorgans to about 30 centiMorgans.

33. (Original) The computer program product of claim 30, wherein said predetermined distance is selected from the range of about 2 centiMorgans to about 15 centiMorgans.

34. (Currently amended) The computer program product of claim 26, each element in said phenotypic data structure ~~representing~~ corresponding to a difference in said phenotype between different strains of said ~~organism~~ species; wherein, for each element in said phenotypic data structure, said different strains of said ~~organism~~ species are selected from said plurality of strains of said ~~organism~~ species represented in said genotypic database.

35. (Currently amended) The computer program product of claim 34, wherein said difference in said phenotype is determined by a measurement of an attribute corresponding to said phenotype in said different strains of said ~~organism~~ species that are represented in said genotypic database.

36. (Currently amended) The computer program product of claim 34, each element in said phenotypic data structure representing a difference in said phenotype between a first cluster of strains of said ~~organism~~ species and a different second cluster of strains of said ~~organism~~ species; wherein, for each element in said phenotypic data structure, said different first and

second cluster of strains of said ~~organism~~ species are selected from a plurality of clusters of strains of said ~~organism~~ species that are represented in said genotypic database.

37. (Currently amended) The computer program product of claim 26, each element in said genotypic data structure representing a variation of at least one component of said locus between different strains of said ~~organism~~ species; wherein, for each element in said genotypic data structure, said different strains of said ~~organism~~ species are selected from said plurality of strains of said ~~organism~~ species represented in said genotypic database.

38. (Currently amended) The computer program product of claim 26, wherein an amount that a variation contributes to said at least one component of said locus between different strains of said ~~organism~~ species is a function of a distance said variation is away from a center of the locus that corresponds to said genotypic data structure.

39. (Original) The computer program product of claim 26, wherein said genotypic data structure represents a plurality of variations that are distributed about the center of said locus, and said instructions for establishing further comprise:

instructions for fitting a distribution of said plurality of variations about the center of said locus with a probability function; and

instructions for weighting each variation by a corresponding value derived from said probability function such that variations further from the center of said locus are downweighted so that they contribute less to said genotypic data structure than loci that are closer to said center of said corresponding locus.

40. (Original) The computer program product of claim 39 wherein said probability function is a Gaussian probability distribution, a Poisson distribution, or a Lorentzian distribution.

41. (Currently amended) The computer program product of claim 26, each element in said genotypic data structure representing a variation of at least one component of said locus between a first cluster of strains of said ~~organism~~ species and a different second cluster of strains of said ~~organism~~ species; wherein, for each element in said genotypic data structure, said different first and second clusters of strains of said ~~organisms~~ species are selected from said plurality of strains of said ~~organism~~ species represented in said genotypic database.

42. (Original) The computer program product of claim 26 wherein said instructions for comparing include instructions for forming said correlation value in accordance with the expression:

$$c(P, G^L) = \frac{\sum^i (p(i) - \langle P \rangle) (g(i) - \langle G^L \rangle)}{\{[\sum^i (p(i) - \langle P \rangle)^2] [\sum^i (g(i) - \langle G^L \rangle)^2]\}^{1/2}}$$

where,

$c(P, G^L)$ is said correlation value;

$p(i)$ is a value of the i^{th} element of said phenotypic data structure;

$g(i)$ is a value of the i^{th} element of said genotypic data structure;

$\langle P \rangle$ is a mean value of all elements in said phenotypic data structure; and

$\langle G^L \rangle$ is a mean value of all elements in said genotypic data structure.

43. (Original) The computer program product of claim 26, wherein said correlation value is weighted by a number of components in said locus.

44. (Original) The computer program product of claim 26, wherein each said component is a single nucleotide polymorphism.

45. (Currently amended) The computer program product of claim 26, wherein said instructions for comparing include instructions for forming said correlation value in accordance with the expression:

$$c(P, G^L) = \frac{[\sum^i (p(i) - \langle P \rangle) (g(i) - \langle G^L \rangle)] \times Z}{\{[\sum^i (p(i) - \langle P \rangle)^2] [\sum^i (g(i) - \langle G^L \rangle)^2]\}^{1/2}}$$

where,

$c(P, G^L)$ is said correlation value;

$p(i)$ is a value of the i^{th} element of said phenotypic data structure;

$g(i)$ is a value of the i^{th} element of said genotypic data structure;

$\langle P \rangle$ is a mean value of all elements in said phenotypic data structure;

$\langle G^L \rangle$ is a mean value of all elements in said genotypic data structure; and

Z is a function of a number of components in said locus having a variation between different ~~strands~~ strains of said ~~organism~~ species.

46. (Original) The computer program product of claim 43, wherein said function is selected from the group consisting of taking the square root of Z, squaring Z, raising Z by the power of a positive integer, taking a logarithm of Z, and taking an exponential of Z.

47. (Original) The computer program product of claim 26, wherein said instructions for comparing include instructions for forming said correlation value in accordance with a correlative measure c_m that is computed in accordance with the expression:

$$c_m(P, G^L) = \frac{[\sum^i (p(i) - \langle P \rangle) (g(i) - \langle G^L \rangle)]}{\{[\sum^i (p(i) - \langle P \rangle)^2]\}^{1/2}}$$

where,

$c_m(P, G^L)$ is said correlative measure;

$p(i)$ is a value of the i^{th} element of said phenotypic data structure;

$g(i)$ is a value of the i^{th} element of said genotypic data structure;

$\langle P \rangle$ is a mean value of all elements in said phenotypic data structure; and

$\langle G^L \rangle$ is a mean value of all elements in said genotypic data structure.

48. (Original) The computer program product of claim 26, wherein said instructions for comparing include instructions for forming said correlation value by an algorithm selected from the group consisting of regression analysis, regression analysis with data transformations, a Pearson correlation, a Spearman rank correlation, a regression tree and concomitant data reduction, partial least squares, and canonical analysis.

49. (Original) The computer program product of claim 26, wherein said instructions for repeating further comprise:

instructions for computing (i) a mean correlation value that represents a mean of each said correlation value formed during instances of said instructions for comparing; and (ii) a standard deviation of said mean correlation value based on each said correlation value formed during instances of said instructions for comparing;

wherein, said one or more genotypic data structures that form a high correlation value relative to all other genotypic data structures compared to said phenotypic data structure by said instructions for comparing are identified by selecting genotypic data structures that form a correlation value that is a predetermined number of standard deviations above said mean correlation value.

50. (Original) The computer program product of claim 26, wherein said genotypic database is a single nucleotide polymorphism database, a microsatellite marker database, a restriction fragment length polymorphism database, a short tandem repeat database, a sequence length polymorphism database, an expression profile database, or a DNA methylation database; and said variation in said genotypic data structure is obtained from said genotypic database.

51. (Cancelled)

52. (Currently amended) A computer system for associating a phenotype with one or more candidate chromosomal regions in a genome of ~~an organism~~ a species, said genome including a plurality of loci, the computer system comprising:

a central processing unit;

a memory, coupled to the central processing unit, the memory storing:

a genotypic database for storing variations in genomic sequences of a plurality of strains of said ~~organism~~ species;

a phenotypic data structure that ~~represents~~ comprises a difference in a phenotype between different strains of said ~~organism~~ species; and

a program module, said program module comprising:

instructions for establishing a genotypic data structure, said genotypic data structure corresponding to a locus ~~selected from a~~ in said plurality of loci, said genotypic data structure ~~representing~~ comprising a variation of at least one component of said locus between different strains of said ~~organism~~ species stored in said genotypic database;

instructions for comparing said phenotypic data structure to said genotypic data structure to form a correlation value; and

instructions for repeating said instructions for establishing and said instructions for comparing, for each locus in said plurality of loci, thereby identifying one or more genotypic data structures that form a high correlation value relative to all other genotypic data structures

that are compared to said phenotypic data structure by said instructions for comparing; wherein the loci that correspond to said one or more genotypic data structures that form a high correlation value represent said one or more candidate chromosomal regions and wherein an amount of said genome that is included in each locus in said plurality of loci is predetermined.

53. (Currently amended) The computer system of claim 52, each element in said phenotypic structure representing a ~~variation~~ difference in ~~said a~~ phenotype between different strains of said ~~organism~~ species; wherein, for each element in said phenotypic data structure, said different strains of said ~~organism~~ species are selected from said plurality of strains of said ~~organism~~ species represented in said genotypic database.

54. (Currently amended) The computer system of claim 53, wherein said difference in a said phenotype is determined by a measurement of an attribute corresponding to said phenotype in said different strains of said ~~organism~~ species that are represented in said genotypic database.

55. (Currently amended) The computer system of claim 52, each element in said phenotypic data structure representing a variation in said phenotype between a first cluster of strains of said ~~organism~~ species and a different second cluster of strains of said ~~organism~~ species; wherein, for each element in said phenotypic data structure, said different first and second cluster of strains of said ~~organism~~ species are selected from a plurality of clusters of strains of said ~~organism~~ species that are represented in said genotypic database.

56. (Currently amended) The computer system of claim 52, each element in said genotypic data structure representing a variation of at least one component of said locus between different strains of said ~~organism~~ species; wherein, for each element in said genotypic data structure, said different strains of said ~~organism~~ species are selected from said plurality of strains of said ~~organism~~ species represented in said genotypic database.

57. (Currently amended) The computer system of claim 52, each element in said genotypic data structure representing a variation of at least one component of said locus between a first cluster of strains of said ~~organism~~ species and a different second cluster of strains of said ~~organism~~ species; wherein, for each element in said genotypic data structure, said different

first and second clusters of strains of said ~~organisms~~ species are selected from said plurality of strains of said ~~organism~~ species represented in said genotypic database.

58. (Original) The computer system of claim 52, wherein said instructions for comparing include instructions for forming said correlation value in accordance with the expression:

$$c(P, G^L) = \frac{\sum^i (p(i) - \langle P \rangle) (g(i) - \langle G^L \rangle)}{\{[\sum^i (p(i) - \langle P \rangle)^2] [\sum^i (g(i) - \langle G^L \rangle)^2]\}^{1/2}}$$

where,

$c(P, G^L)$ is said correlation value;

$p(i)$ is a value of the i^{th} element of said phenotypic data structure;

$g(i)$ is a value of the i^{th} element of said genotypic data structure;

$\langle P \rangle$ is a mean value of all elements in said phenotypic data structure; and

$\langle G^L \rangle$ is a mean value of all elements in said genotypic data structure.

59. (Original) The computer system of claim 52, wherein said instructions for comparing include instructions for forming said correlation value by an algorithm selected from the group consisting of regression analysis, regression analysis with data transformations, a Pearson correlation, a Spearman rank correlation, a regression tree and concomitant data reduction, partial least squares, and canonical analysis.

60. (Original) The computer system of claim 52, wherein said instructions for repeating further comprise:

instructions for computing (i) a mean correlation value that represents a mean of each said correlation value formed during instances of said instructions for comparing; and (ii) a standard deviation of said mean correlation value based on each said correlation value formed during instances of said instructions for comparing;

wherein, said one or more genotypic data structures that form a high correlation value relative to all other genotypic data structures compared to said phenotypic data structure by said instructions for comparing are identified by selecting genotypic data structures that form a correlation value that is a predetermined number of standard deviations above said mean correlation value.

61. (Original) The computer system of claim 52, wherein said genotypic database is a single nucleotide polymorphism database, a microsatellite marker database, a restriction fragment length polymorphism database, a short tandem repeat database, a sequence length polymorphism database, an expression profile database, or a DNA methylation database; and said variation in said genotypic data structure is obtained from said genotypic database.

62-70. (Cancelled)

71. (Currently amended) A computer program product for use in conjunction with a computer system, the computer program product comprising a computer readable storage medium and a computer program mechanism embedded therein, the computer program mechanism comprising:

a program module for associating a phenotype with one or more candidate chromosomal regions in a genome of said ~~organism~~ species, said genome including a plurality of loci, said program module comprising:

instructions for accessing a genotypic data structure, said genotypic data structure corresponding to a locus ~~selected from a~~ in said plurality of loci, said genotypic data structure ~~representing~~ comprising a variation of at least one component of said locus between different strains of said ~~organism~~ species stored in a genotypic database;

instructions for comparing a phenotypic data structure to said genotypic data structure to form a correlation value wherein said phenotypic data structure comprises a difference in said phenotype between different strains of said species; and

instructions for repeating said instructions for establishing and instructions for comparing for each locus in said plurality of loci, thereby identifying one or more genotypic data structures that form a high correlation value relative to all other genotypic data structures that are compared to said phenotypic data structure by said instructions for comparing; wherein the loci that correspond to said one or more genotypic data structures that form a high correlation value represent said one or more candidate chromosomal regions.

72. (New) The computer system of claim 52, wherein said amount is selected from a value in the range of about 0.01 centiMorgans to about 100 centiMorgans.

73. (New) The computer system of claim 52, wherein said amount is selected from a value in the range of about 5 cM to about 30 cM.

74. (New) The computer system of claim 52, wherein an instance of said instructions for establishing comprise selecting a locus that is centered on a portion of said genome that is a predetermined distance away from the locus that was selected by a previous instance of said instructions for establishing.

75. (New) The computer system of claim 74, wherein said predetermined distance is measured in centiMorgans.

76. (New) The computer system of claim 74, wherein said predetermined distance is selected from the range of about 0.0001 centiMorgans to about 30 centiMorgans.

77. (New) The computer system of claim 74, wherein said predetermined distance is selected from the range of about 2 centiMorgans to about 15 centiMorgans.